

University of Groningen

Perceptual interactions in human vision and implications for information visualization

van den Berg, Ronald

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2009

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

van den Berg, R. (2009). *Perceptual interactions in human vision and implications for information visualization*. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 3

Perceptual Dependencies in Information Visualization Assessed by Complex Visual Search

This chapter has been published as:

Ronald van den Berg, Frans W. Cornelissen and Jos B. T. M. Roerdink. *Perceptual Dependencies in Information Visualization Assessed by Complex Visual Search*. **ACM Transactions on Applied Perception**, 4(4):1-21, 2008.

Abstract

A common approach for visualising data sets is to map them to images in which distinct data dimensions are mapped to distinct visual features, such as colour, size and orientation. Here, we consider visualisations in which different data dimensions should receive equal weight and attention. Many of the end user tasks performed on these images involve a form of visual search. Often, it is simply assumed that features can be judged independently of each other in such tasks. However, there is evidence for perceptual dependencies when simultaneously presenting multiple features. Such dependencies could potentially affect information visualisations that contain combinations of features for encoding information, and thereby bias subjects into unequally weighting the relevance of different data dimensions. We experimentally assess 1) the presence of judgment dependencies in a visualisation task (searching for a target node in a node-link diagram) and 2) how feature contrast relates to salience. From a visualisation point of view, our most relevant findings are that (a) to equalise saliency (and thus bottom-up weighting) of size and colour, colour contrasts have to become very low. Moreover, orientation is less suitable for representing information that consists of a large range of data values because it does not show a clear relationship between contrast and salience; (b) colour and size are features that can be used independently to represent information, at least for the range of colours that were used in our study; (c) the concept of (static) feature salience hierarchies is wrong; how salient a feature is compared to another is not fixed, but a function of feature contrasts; (d) final decisions appear to be as good an indicator of perceptual performance as indicators based on measures obtained from individual fixations. Eye-tracking therefore does not necessarily present a benefit for user studies that aim at evaluating performance in search tasks.

3.1 Introduction

Information visualisation helps to locate and understand patterns and relationships in large data sets by transforming them into sensory images. While a given data set can be mapped to an image in nearly infinite ways, not every mapping will be equally effective. One of the goals of information visualisation research is to reveal the principles that determine whether a particular mapping is effective. As these principles depend on how the human brain processes and interacts with visual information, it has been recognised that the study of perceptual aspects should play a prominent role in visualisation practices [Bertin, 1983; Ebert, 2004; House & Ware, 2002; Interrante, 2004; Johnson, 2004; Tory & Müller, 2004; Ware, 2000].

A general approach for visualising data sets is to map distinct data dimensions to distinct visual object features (such as colour, texture, size, shape and orientation). In line with traditional theories of visual search (e.g., Desimone & Duncan [1995]; Treisman & Gelade [1980]) it is then often assumed that such features are judged independently of each other, in the sense that appearance of the one does not depend on that of the other(s).

There is, however, experimental evidence challenging this view of independent feature processing. In particular, colour appears to be a very dominant feature. In previous experiments we found that judgment of orientation and size in a basic search task was contingent on whether one simultaneously searches for colour, while judgment of colour did not show the reverse dependency [Hannus *et al.*, 2006] (some details about this experiment will be provided in Section 3.2.4). Nothdurft 2000 found that adding colour to an oriented visual object increases its visual salience more than adding orientation to a coloured object. It has also been found that search for colour is faster than search for shape [Luria & Strauss, 1975]. Furthermore, Callaghan 1989 reported that judgment of shape-defined boundaries is affected by hue variation but that the reverse is not true. A similar effect was reported by Snowden 1998, who found that irrelevant variations in colour, depth, and combinations of colour and depth produce detrimental effects in performance on texture segregation tasks. Williams and Reingold 2001 found that subjects are more likely to fixate non-target distractors that have the same colour as the target than distractors that have the same shape and/or orientation. In addition to these psychophysical findings, there is a substantial amount of physiological evidence that indicates that features are multiplexed (e.g., chromatically tuned orientation selective cells), at least at the early stages of processing by strictly separated brain regions/cells [Gegenfurtner, 2003; Roe & Ts'o, 1999; von der Heydt *et al.*, 2003; Yoshioka & Dow, 1996]. Even though it is not clear how this relates to the more feature-specific processing that is assumed to be present in later stages of visual perception [Livingstone & Hubel, 1988; Zeki & Shipp, 1988], it has been suggested that processing at these early stages may determine feature salience and as such search performance [Li, 2002].

Such perceptual feature dependencies may have implications for information

visualisation design. For example, information that is visualised by orientation or size could become less salient (and therefore more difficult to judge) when colour-coded information is added or when colour contrast is being changed. Effects such as these will make it hard for the visualisation designer to predict and control the salience of displayed information and should therefore be avoided as much as possible.

Because of differences in complexity and duration of tasks and stimuli, one should refrain from straightforward generalisation of findings from basic psychophysical experiments to the domain of information visualisation applications. Low-level effects found by psychophysical experiments might for example be negligible in the problem-solving strategies that are used in information visualisation tasks. The above cited findings should therefore be reassessed in a visualisation context.

Currently, the issue of feature interactions is not well understood [Ebert, 2004] and remains unexplored by most visualisation scientists. One exception to this concerns the work of Healey and colleagues, who have conducted psychophysical experiments to test for possible perceptual interactions between the features they visualise their data with. For example, Healey and Enns 1998 used psychophysical methods to test for interference between the dimensions (texture height, density and regularity) of the perceptual texture elements ('pexels') that were used to visualise their data. They found that height, regularity and density of background pexels interfered with short, dense and sparse pexels. In a later study it was assessed whether colour could be added to pexels without interfering with one of the other pexel dimensions [Healey & Enns, 1999]. It was found that variations of height and density had no effect on colour segmentation, whereas random colour patterns interfered with texture segmentation. Discriminability of features was not matched in these studies, however. As a result, it is possible that, for example, variations in height and density were of different (perceptual) magnitude than the variations in colour. If this was indeed the case, then the interference effect might be explained by a design asymmetry and could possibly be removed by reducing colour contrast (or increasing height and density contrasts).

Here, our primary question is whether previously found feature judgment interactions have potential relevance for information visualisation. We experimentally assessed this using a relatively complex and visualisation-realistic visual search task (many objects, a large range of feature values, and relatively long task duration). As colour, size, and - to a lesser extent - orientation are frequently used features in information visualisation and most evidence points towards colour as a potential interfering feature, the experiments were carried out with combinations of colour and size and combinations of colour and orientation. Prior to the experiments we matched discriminability of the features in order to avoid design asymmetries as well as bias of the subjects' attention towards a feature with higher salience than the others.

3.2 Methods

3.2.1 Subjects

Six subjects participated in the colour/size experiment (three females and three males, one of them author RB). Four of these subjects also participated in the colour/orientation experiment. All participants had normal or corrected-to-normal visual acuity and normal colour vision.

3.2.2 Apparatus

Stimuli in the form of node-link diagrams were presented on a 22-inch monitor at a resolution of 2048x1536 pixels and with a refresh rate of 75 Hz. For display of the diagrams we made use of the force-directed graph layout algorithm of Cytoscape (<http://www.cytoscape.org>; [Shannon *et al.*, 2003]). Stimulus presentation and data collection were done using Matlab in combination with the Psychophysics and Eyelink Toolbox extensions [Brainard, 1997; Cornelissen *et al.*, 2002; Pelli, 1997]. Eye movements were recorded with an Eyelink II eye tracker (SR Research, Ltd., Mississauga, Ontario, Canada) with a temporal frequency of 250 Hz. Subjects viewed the stimuli at a distance of about 45 cm. A chin-rest assisted them in reducing head movements as much as possible.

3.2.3 Stimuli

Conjunction search stimuli

The stimuli consisted of a cue followed by a node-link diagram made up of 50 nodes and 70 (task-irrelevant) edges (Figure 3.1). Nodes were either discs with a particular colour and size (colour/size conditions) or bars with a particular colour and orientation (colour/orientation conditions). In each trial, one of the nodes was randomly chosen to be the target and was assigned a random colour and size or orientation (for details about the colours, sizes and orientations used, see below). The other 49 nodes were distractors and were also assigned a random colour and size or orientation, with the restriction not to be identical to the target (and thus a distractor could, for example, have a different size but the same colour as the target). The cue was identical to the target and was shown in the centre of an otherwise blank screen for two seconds. The target and distractors had a luminance of approximately 9.3 cd/m^2 and were viewed against a grey background (approximately 7.1 cd/m^2).

Single feature search stimuli

Stimuli for single feature search conditions were exactly the same as the conjunction search stimuli, except for the cues. In single feature search conditions, the

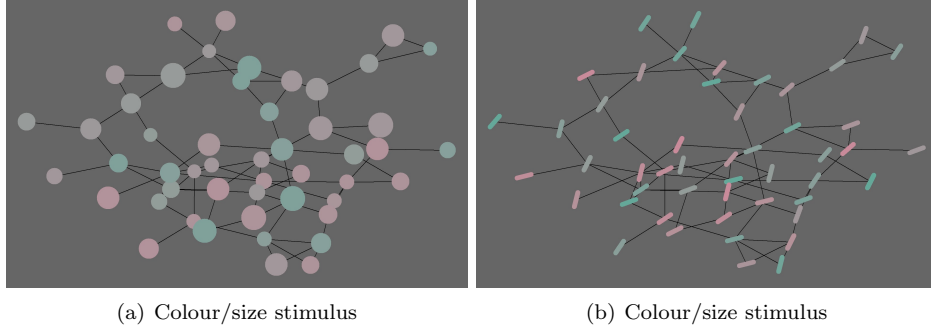


Figure 3.1: *Example stimuli*

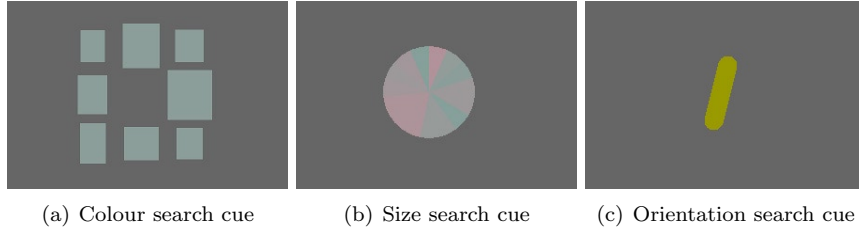


Figure 3.2: *Cues in single feature search conditions. Cues in conjunction search conditions were identical to the target*

cues contained only information about the feature to be searched. Size information was removed by using cues as displayed in Figure 3.2(a); colour information was removed by using cues as displayed in Figures 3.2(b) and 3.2(c).

3.2.4 Definition and matching of contrasts

Colour contrasts were created by increasing (decreasing) the output luminance of the red monitor channel and simultaneously decreasing (increasing) the output luminance of the green channel with respect to a base colour. This was done in such a way that luminance was held constant across search items. CIE-coordinates of the default range (see below) ranged from ($x=.271$, $y=.311$; green) to ($x=.286$, $y=.305$; red). Size contrasts were created by increasing or decreasing the diameter of the nodes with respect to the base size (approximately 0.9° of visual angle). Orientation contrasts were created by tilting nodes in clockwise or counter-clockwise direction with respect to the base orientation (45° with respect to horizontal direction). To avoid design asymmetries or subjects being biased towards one feature or another due to large salience differences, we generated perceptually matched colour, size and orientation ranges prior to the experiments. For this, we first determined perceptually matched step-sizes for the three fea-

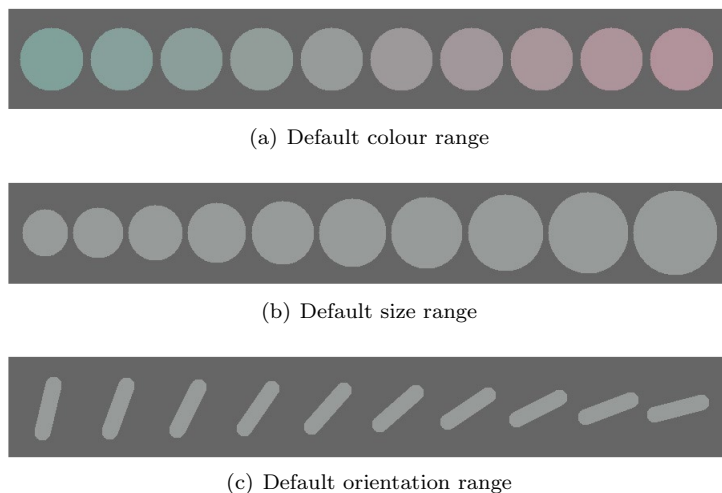


Figure 3.3: *Default feature ranges used in the experiments. The perceptual distance (the deltas) between each two consecutive colours, sizes, and orientations was fixed. Also, the perceptual distance between two consecutive colours was matched with respect to the distance between two consecutive sizes and orientation (for an explanation of the low colour contrast, see Section 3.4.1)*

tures (we will refer to these as delta values Δc , Δs , and Δo , respectively). Using these deltas, for each condition a range of ten different values was created for each feature (Figure 3.3).

This ensured that the perceptual distance (and thus, discriminability) between two consecutive values was the same for all features. Matches were established based on data from previous experiments [Hannus *et al.*, 2006]. In those experiments subjects performed a large amount of single feature search trials (for colour, orientation and size) with 10 different target-distractor contrasts. From the resulting data we computed for all three features the mean psychometric curve (sigmoid fits), describing the relationship between feature contrast and search performance. We then defined colour, size, and orientation deltas (Δc , Δs , and Δo) as the difference between the contrasts needed for a performance of 70% and 50%, respectively (note that due to the linearity of most of the psychometric curve, we could just as well have used other performance values; e.g., 30-50% or 40-60%). This approach is comparable to the use of just-noticeable differences (JNDs), but other methods for balancing feature saliency are conceivable as well (e.g., Nothdurft [2000]).

To demonstrate the matching method we will show how it worked for size (colour and orientation contrasts were generated analogously). The mean psychometric curve for size discrimination is shown in Figure 3.4. It appeared that a \log_{10} difference of 1.24 (17.4%) between target and distractor diameter was needed

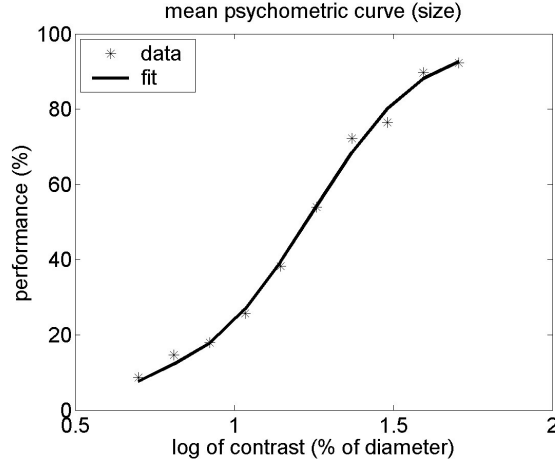


Figure 3.4: Mean psychometric curve for size. Step-size Δs was defined as the difference between the contrasts resulting in 70% and 50% performance accuracy.

to obtain a performance of 50% accuracy and a \log_{10} difference of 1.38 (24.0%) was needed for a performance of 70%. The size delta was thus set to 6.6%. Colour and orientation deltas were determined in a similar way ($\Delta c = 1.2\%$, $\Delta o = 7^\circ$).

In all experimental conditions, feature dimensions contained 10 different values. In conditions where contrast was set to default, a range with 10 values was created by modulating the feature's base value with the delta value. This was computed in the case for size as follows: $size_i = base_size + (i - 5\frac{1}{2}) * \Delta s$, $i = 1, 2, \dots, 10$.

In this range of sizes, the distance between each two consecutive sizes is fixed and it is therefore expected that discriminability of sizes 1 and 2 is the same as that of sizes 2 and 3, etc. Moreover, due to the matching, discriminability of two consecutive sizes is expected to be the same as that of two consecutive (default) colours and orientations. As a result, equal search performance is expected in single feature search tasks when using the default ranges.

One of the experimental questions was whether changing contrast in one feature dimension interferes with search performance in other feature dimensions. Therefore, in addition to conditions with default feature ranges, we included conditions in which contrast of one feature was reduced or enhanced while that of the other was kept at default. To achieve this, contrast modification factors C, S, and O were added to the feature range equations. For size we thus get: $size_{s,i} = base_size + (i - 5\frac{1}{2}) * S * \Delta s$, $i = 1, 2, \dots, 10$.

A contrast modification factor set to 1 resulted in stimuli with default contrast, while modification factors smaller than 1 reduced the contrast and modification factors larger than 1 enhanced it; 0 resulted in uniform ranges (base values). Since conditions only differed from each other with respect to search type (single feature,

conjunction) and contrast modification factor, in the remainder of this paper we will refer to them by their contrast modifications (e.g., 'conjunction search with $C=S=1$ ' is conjunction search with matched contrasts and 'size search with $S=1$ and $C=0$ ' is single feature search of size with no variation in the colour dimension).

The range of sizes that we could use for the experiment was constrained by some practical aspects. Node sizes too large would result in excessive overlap of nodes, while node sizes too small would make it impossible to generate small gradual increments (node diameters could only be a discrete number of pixels). We chose a size range that avoided both problems and then matched the colour and orientation ranges to this.

3.2.5 Procedure

Participants were instructed to search for the target node in the node-link diagram. They were informed about the identity of the target by means of a cue, as described in Section 3.2.3. They were asked to keep fixating at the selected node until the end of the trial in case they thought to have found the target. Eye movements were recorded during display of the node-link diagram, which were shown for 4.5 seconds. Search time was limited for two reasons: (i) to ensure that all subjects would opt for speed in the inevitable speed-accuracy trade-off that has to be made in search tasks like these and (ii) to limit total experiment time. After each trial, feedback was given by highlighting the target node and the one selected by the subject (i.e., the one last fixated).

The task was performed under eight different conditions with colour/size stimuli and under nine different conditions with colour/orientation stimuli. As mentioned earlier, conditions only differed from each other with respect to search type (single feature, conjunction) and contrast modification factors C , S , and O . There were four colour/size conjunction search conditions: one with matched contrasts ($C=S=1$), one with reduced colour contrast ($C=0.5$), and two with enhanced size contrast ($S=1.5$, $S=2$). There were also four colour/size single feature search conditions: colour search with ($S=1$) and without ($S=0$) task-irrelevant size contrast, and size search with ($C=1$) and without ($C=0$) task-irrelevant colour contrast. When there was no contrast in a feature dimension, a random value from the default feature range was chosen and assigned to all nodes.

The experiment contained five colour/orientation conjunction search conditions: one with matched contrasts ($C=O=1$), three with modified colour contrast ($C=0.5$, 2, 4), and one with doubled orientation contrast ($O=2$). The four single feature search conditions were analogous to those with colour/size stimuli.

Some asymmetries can be observed across the conditions. The reasons are as follows. Earlier we observed that colour tends to dominate in conjunction search when having matched salience for single feature search. Our intention to assess whether this colour dominance effect can be undone by either decreasing colour contrast or increasing size contrast explains the asymmetry in the design of the first experiment. Orientation contrast is limited to a step-size that does not

result in a feature range exceeding 180° and could therefore not be increased much further (with $\Delta o = 7$, this range would already be exceeded when using $O=3$). We used $C=0.5, 1, 2, 4$ with the intention to get a more accurate picture of the relationship between (colour) contrast and feature salience; $C=4$ is the maximum, because higher contrasts would exceed the DAC range of our monitor.

Prior to the experiments, 60 random node-link diagrams were generated (as the one shown in Figure 3.1). Each of these networks appeared exactly once in each condition, in a fixed order. Colour/size and colour/orientation conditions were measured in two separate sessions. Within these sessions, conditions were presented in separate, randomly ordered blocks of 60 trials. At the start of each block, a calibration procedure was performed.

3.2.6 Analysis

Relying on the assumption that a fixation was always made to inspect the node nearest to the point of fixation, we transformed the sequences of recorded fixations into sequences of node inspections by looking up for each fixation what node was closest to it. Based on these node inspection sequences we computed for every inspection the error for each of the relevant features. This error was defined as the difference between the feature value index of the target (a number between 1 and 10; see Section 3.2.4) and that of the inspected node. All statistical tests were carried out on these errors.

Since it is possible that different strategies are used in the search stage and the making of the final decision at the end of the trial, we distinguish between 'search fixations' and 'decision fixations' in the analysis. The final decision errors can be seen as a measure of eventual task performance and are, therefore, the most important ones from a visualisation standpoint. The search stage errors can provide insight in the search strategies and possibly explain the final decision results. In case the previously found feature judgment interactions are merely low-level effects, we can expect to find them in the search stage while they do not exist in the final decisions.

Analysis of final decisions includes only participants' eventual choices at the end of the trials. As it often occurred that a subject was still in the process of searching at the end of a trial, we could not simply use all final fixations for this analysis. As a criterion to filter out trials in which subjects were still searching when the trial ended, we only included those in which the last fixation had a duration that was at least two standard deviations longer than the mean duration of all other fixations in that trial. The second type of analysis assesses the error during search. It is likely that not only the very last fixation of a trial was directly related to a subject's decision, but also a couple of fixations preceding the last one (correction saccades). Since such fixations are not directly related to the search process, we excluded them from the analysis. We chose to be conservative and omitted the last 25% of the fixations of every trial. Analyses of these data were done on the means of the signals.

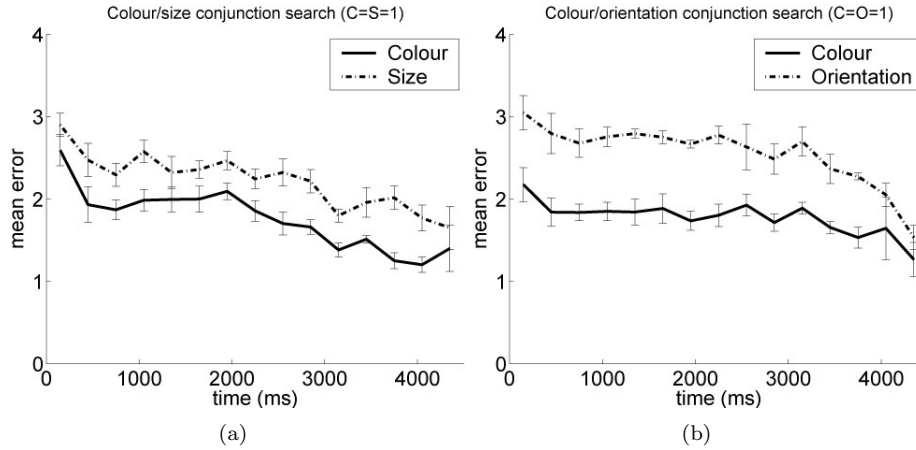


Figure 3.5: Time courses of errors in (a) colour/size and (b) colour/orientation search with matched contrasts. Trials were split into 15 intervals of 300 ms and mean error is shown for each interval. Note the high correlations between the signals (R^2 is .86 for colour/size and .85 for colour/orientation). Bars represent standard errors.

All statistical analyses consisted of repeated measures analysis of variances tests (ANOVA), with a significance level of 0.05. We tested whether (i) in single feature search, variation in a task-irrelevant feature interferes with search performance, (ii) in conjunction search, contrast of one feature affects search performance of the other, and (iii) there was an interaction effect between search type (single feature, conjunction) and feature (colour, orientation/size).

3.3 Results

In the following section we subsequently present results regarding the error signals, salience matching, the colour/size experiment and the colour/orientation experiment.

3.3.1 Error signals

As a first visual assessment of the data, we inspected the error signals before beginning the statistical analysis. Figure 3.5 shows the mean error signals in colour/size conjunction search with $C=S=1$ (left) and colour/orientation search with $C=O=1$ (right). The other conditions resulted in similar signals. Please note that these time courses are only meant as a first (informal) assessment of the data and that all statistical analyses below have been performed on mean errors, as described in the previous section.

| | Conjunction Search | | | | Single Feature Search | | Interaction effect |
|--------------------|---------------------------------|------------|--------------------------------|------------|---|---|-------------------------------------|
| | Effect of colour contrast on... | | Effect of size contrast on ... | | Effect of size contrast on colour error | Effect of colour contrast on size error | <i>Search type</i> × <i>feature</i> |
| | colour error | size error | colour error | size error | | | |
| Decision fixations | .042 | .22 | .83 | .0013 | .18 | .37 | .33 |
| Search fixations | .003 | .48 | .57 | <<.01 | .32 | .36 | .07 |

Table 3.1: Summary of the statistics regarding the studied effects in the colour/size data (*p-values*).

A first observation is that colour error is smaller than both the size and orientation error. This might make one wonder whether contrasts were matched correctly. However, these are error signals from conjunction search conditions, while contrasts were matched for single feature search. Therefore, the quality of matching will be assessed based on the single feature search data (see next section). A second observation concerns the overall shape of the signals. We see a relatively large decrease in the first 500-1000 ms in all signals, followed by a slow decrease over the rest of the time course. Regarding the orientation error we see that there is again a quick decrease at the end of the trials. A final observation is the high correlation between the error signals, suggesting that errors were minimised for both features in parallel.

3.3.2 Quality of contrast matches

To assess the quality of the contrast matches, we checked whether there were any significant differences between colour and size, and colour and orientation errors in single feature search conditions with irrelevant contrast in the second feature. The difference between colour and size error in these conditions was not significant either for decision fixations [$F(1, 5) = 2.98$, $p = .14$] or search fixations [$F(1, 5) = 5.89$, $p = .060$], indicating that the task difficulty (and thus feature discriminability) was comparable for these features. The difference between colour and orientation error during search was significant for search fixations [$F(1, 3) = 38.4$, $p = .009$] but not for decision fixations [$F(1, 3) = .606$, $p = .49$]. It thus appears that despite the matching procedure, orientation search was more difficult than colour search.

3.3.3 Colour/size data

This section gives a detailed description of the statistical analyses for all effects that have been studied in the colour/size data. A summary (*p-values*) can be found in Table 3.1.

Decision fixation errors (colour/size)

Of all colour/size trials, 68% met the criterion that the duration of the last fixation was at least two standard deviations longer than the mean fixation duration (see Section 3.2.6 for the rationale behind this criterion). Repeated measures one-way analyses of variance (ANOVA) reveal that in conjunction search (Figure 3.6, top) there is a significant effect of colour contrast on colour error [$F(1, 5) = 7.37$, $p = .042$] but not on size error [$F(1, 5) = 1.94$, $p = .22$]. There is also a significant effect of size contrast on size error [$F(2, 10) = 13.81$, $p = .0013$] but not on colour error [$F(2, 10) = .19$, $p = .83$]. This indicates that in conjunction search, colour and size contrasts determine decision performance with respect to the feature itself but not to the other.

In single feature search (Figure 3.6, bottom) there is no significant difference between error in colour search with and without size contrast [$F(1, 5) = 2.39$, $p = .18$]. The same holds for size error in single feature search with and without (task-irrelevant) colour contrast [$F(1, 5) = .99$, $p = .37$]. It thus seems that colour search is not affected by task-irrelevant variation in the size dimension and vice versa.

No significant interaction effect between factors search type (conjunction, single feature) and feature (colour, size) was found [$F(1, 5) = 1.18$, $p = .33$] (Figure 3.7). This was tested by using the data from the conjunction search conditions with $C=S=1$ (Figure 3.6, top), those of the colour search condition with irrelevant size contrast ($C=S=1$; Figure 3.6(c)), and those of the size search condition with irrelevant colour contrast ($C=S=1$; Figure 3.6(d)). This indicates that the difference between colour error in single feature search and conjunction search is similar to that between size error in single feature and conjunction search.

Error during search (colour/size)

In conjunction search (Figure 3.8, top) there is a significant effect of colour contrast on colour error [$F(1, 5) = 30.0$, $p = .003$] but not on size error [$F(1, 5) = .57$, $p = .48$]. There is also a significant effect of size contrast on size error [$F(2, 10) = 29.3$, $p < .0001$] but not on colour error [$F(2, 10) = .59$, $p = .57$]. In single feature search (Figure 3.8, bottom) there is no significant difference between error in colour search with and without (task-irrelevant) size contrast [$F(1, 5) = 1.20$, $p = .32$]. The same holds for size error in single feature search with and without (task-irrelevant) colour contrast [$F(1, 5) = 1.00$, $p = .36$]. No significant interaction effect between search type and feature was found [$F(1, 5) = 5.31$, $p = .070$] (Figure 3.9).

In summary, we see exactly the same pattern of results for search fixation errors as we saw for the decision fixation errors: (i) in conjunction search, feature performance is determined by its contrast but not by the contrast of the other feature, (ii) in single feature search, task-irrelevant variation in the second feature

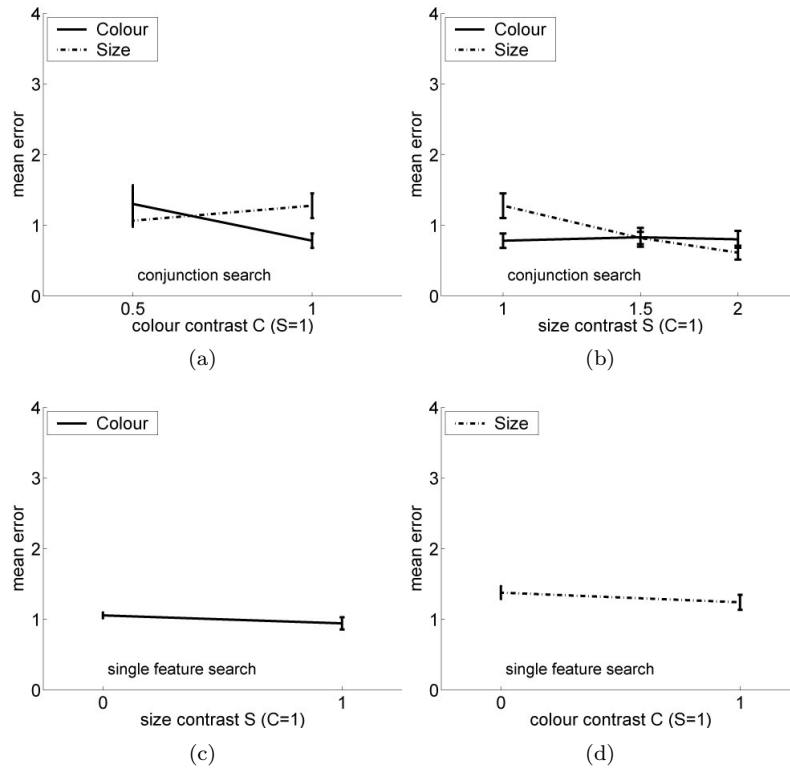


Figure 3.6: Mean errors for decision fixations in conjunction search (top) and single feature search (bottom) conditions, $N=6$. Bars represent standard errors; x-axis is in log scale.

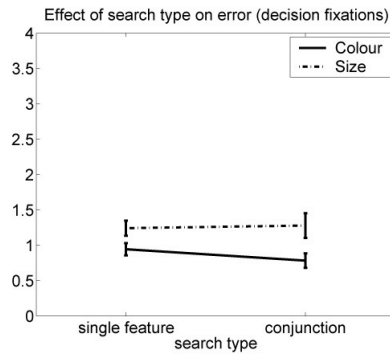


Figure 3.7: Mean error of decision fixations in single feature and conjunction search, $N=6$. All data shown is from search tasks with matched contrasts ($C=S=1$). Bars represent standard error.

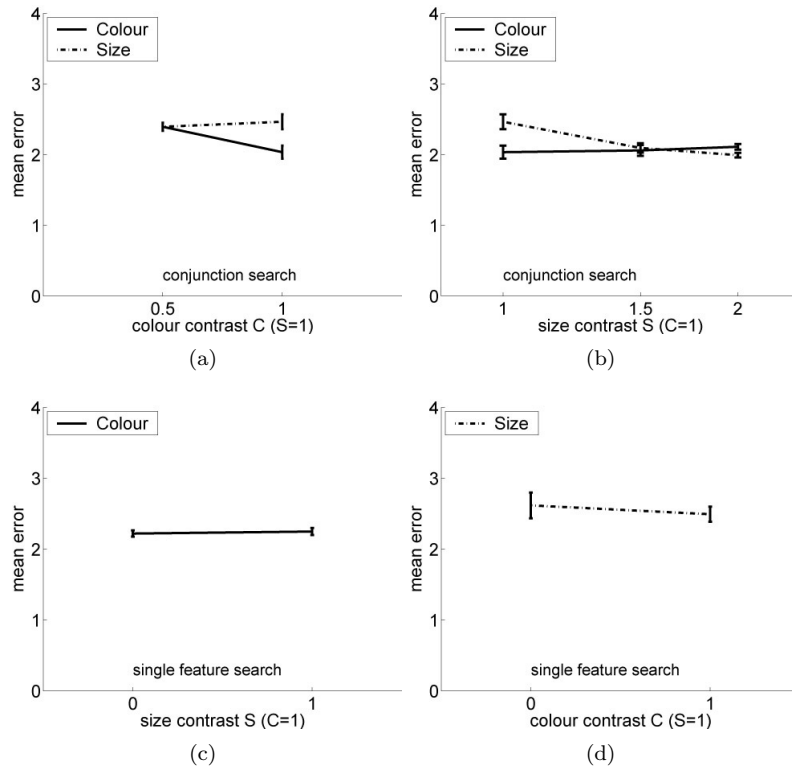


Figure 3.8: Mean errors during search in conjunction search (top) and single feature search (bottom) conditions, $N=6$. Bars represent standard errors; x -axis is in log scale.

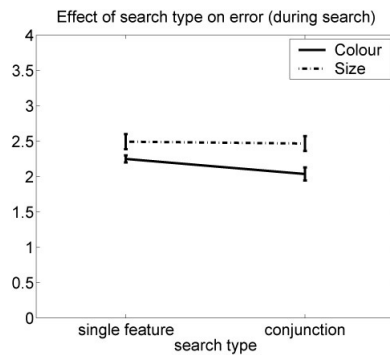


Figure 3.9: Error of search stage fixations in single feature and conjunction search, $N=6$. All data shown is from search tasks with matched contrasts ($C=S=1$). Bars represent standard error.

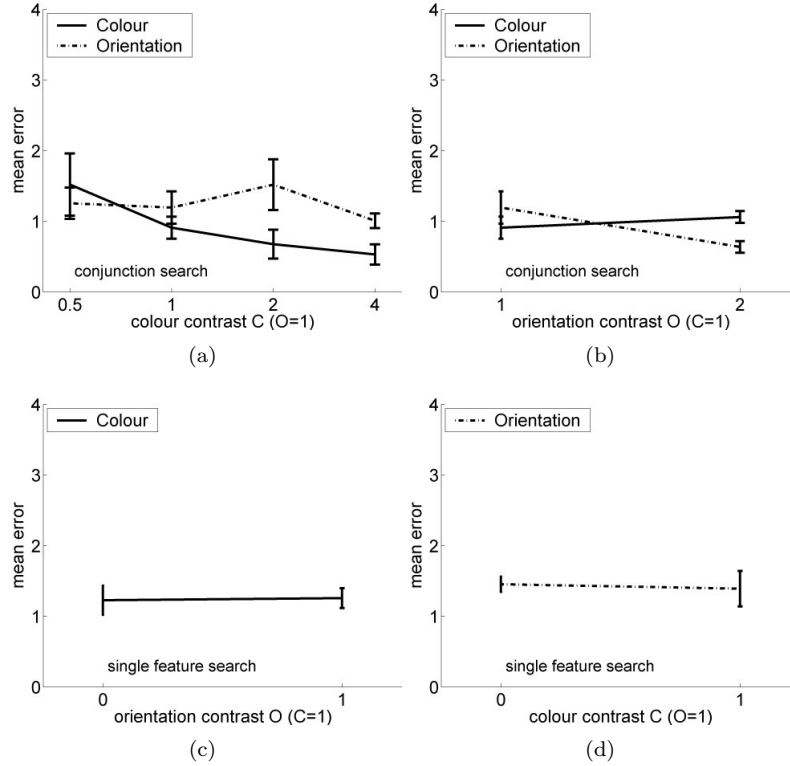


Figure 3.10: Mean errors for decision fixations in conjunction search (top) and single feature search (bottom) conditions, $N=4$. Bars represent standard errors.

dimension does not affect performance, and (iii) there is no interaction of factors search type and feature on performance.

3.3.4 Colour/orientation data

This section gives a detailed description of the statistical analyses of all effects that have been studied in the colour/orientation data. A summary (p-values) can be found in Table 3.2.

Decision fixation errors (colour/orientation)

A total of 56% of the trials from the colour/orientation experiment met the criterion that the duration of the last fixation was at least two standard deviations longer than the mean fixation duration.

In conjunction search (Figure 3.10, top) there is a significant effect of colour contrast on colour error [$F(3, 9) = 4.28$, $p = .039$] but not on orientation error

| | Conjunction Search | | | | Single Feature Search | | Interaction effect |
|--------------------|---------------------------------|-----------|-------------------------------|-----------|--|--|-----------------------|
| | Effect of colour contrast on... | | Effect of ort contrast on ... | | Effect of ort contrast on colour error | Effect of colour contrast on ort error | Search type x feature |
| | colour error | ort error | colour error | ort error | | | |
| Decision fixations | .039 | .13 | .48 | .062 | .74 | .70 | .49 |
| Search fixations | <<.01 | .005 | .03 | .95 | .86 | .09 | .45 |

Table 3.2: Summary of the statistics regarding the studied effects in the colour/orientation data (*p*-values).

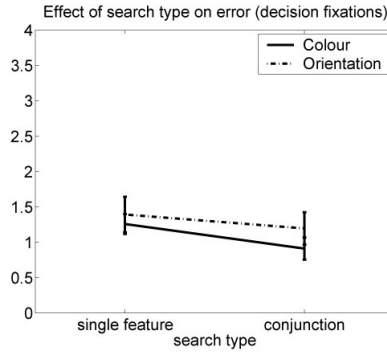


Figure 3.11: Mean error of decision fixations in single feature and conjunction search, $N=4$. All data shown is from search tasks with matched contrasts ($C=O=1$). Bars represent standard error.

$[F(3, 9) = 2.424, p = .13]$ Orientation contrast has no significant effect on orientation error $[F(1, 3) = 8.50, p = .062]$ or colour error $[F(1, 3) = .65, p = .48]$. Thus it seems that colour contrast determines colour but not orientation error, while orientation contrast affects neither of them.

In single feature search (Figure 3.10, bottom) no significant difference between colour error in single feature search with and without (task-irrelevant) orientation contrast was found $[F(1, 3) = .13, p = .74]$. The same is true for orientation error in single feature search with and without (task-irrelevant) contrast in colour $[F(1, 3) = .18, p = .70]$. So, we again see that variation in the task-irrelevant second feature dimension does not affect single feature search performance.

Again, no interaction effect between search type and feature was found $[F(1, 3) = .63, p = .49]$ (Figure 3.11).

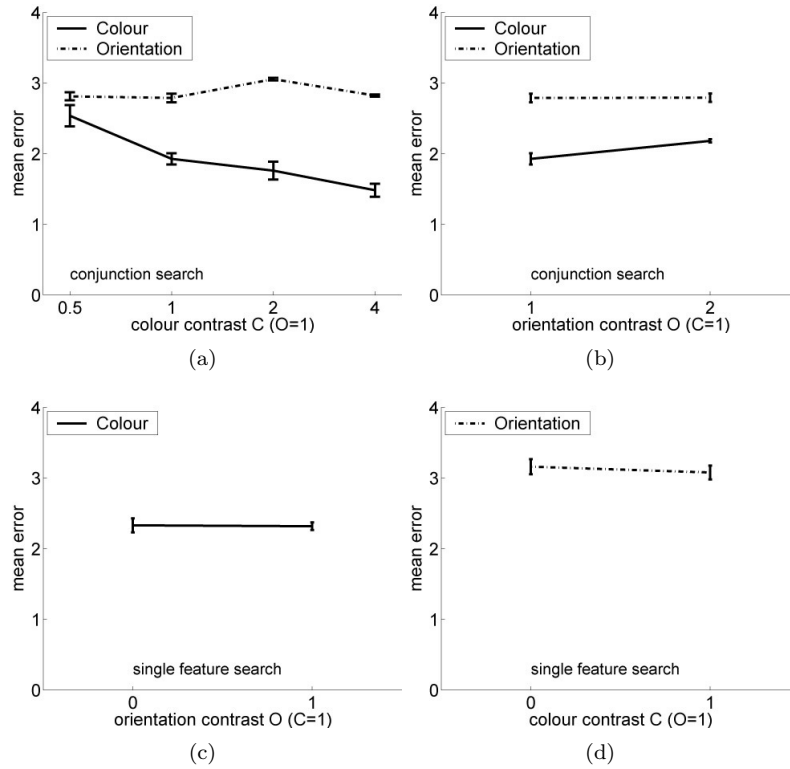


Figure 3.12: Mean errors during search in conjunction search (top) and single feature search (bottom) conditions, $N=6$. Bars represent standard errors.

Error during search (colour/orientation)

In conjunction search (Figure 3.12, top) there is a significant effect of colour contrast on colour error [$F(3, 9) = 32.4$, $p < .0001$] and also on orientation error [$F(3, 9) = 8.82$, $p = .005$]. Also, there is a significant effect of orientation contrast on colour error [$F(1, 3) = 16.2$, $p = .03$] but not on orientation error [$F(1, 3) = .004$, $p = .95$]. It appears that colour error depends on both colour and orientation contrast, whereas orientation error depends on colour contrast only.

In single feature search (Figure 3.12, bottom) no significant difference between colour error in single feature search with and without (task-irrelevant) orientation contrast was found [$F(1, 3) = .036$, $p = .86$]. Also the difference between orientation error in single feature search with and without (task-irrelevant) colour contrast is not significant [$F(1, 3) = 6.34$, $p = .09$].

No interaction effect between search type and feature was found [$F(1, 3) = .76$, $p = .45$] (Figure 3.13).

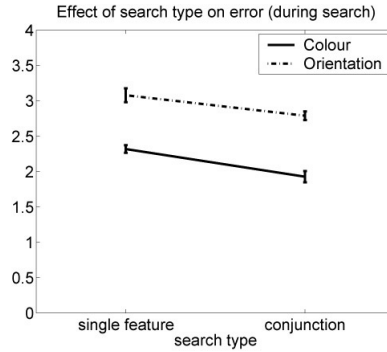


Figure 3.13: Mean error of search stage fixations in single feature and conjunction search, $N = 4$. All data shown is from search tasks with matched contrasts ($C=O=1$). Bars represent standard error.

3.4 Discussion

The primary goal of our experiments was to determine whether earlier reported feature judgment interactions have relevance for information visualisation. We will first discuss colour/size and colour/orientation search interactions in the light of our results and then consider some secondary findings. Thereafter, we will discuss some more general considerations. We will start with briefly discussing the saliency matching procedure.

3.4.1 Saliency matching

As mentioned previously, in the present experiments, we consider visualisations in which different data dimensions should receive equal weight and attention. To avoid any strong *a priori* bottom-up attentional biases for one or the other feature in the displays in our experiments, we attempted to match feature saliency in our stimuli. We did so by choosing feature contrasts that resulted in equal performance improvement in simple visual search tasks.

A consequence of this requirement (and practical restrictions related to display size and resolution) is that colour contrasts became relatively weak. (Another way to look at this is that colour, when used at higher contrasts, is an extremely powerful attentional cue). Hence, a first message from this study is that when equal perceptual weighting of data dimensions coded using size and colour features is required, very low colour contrasts have to be used. In Section 3.4.4, we will further discuss the suitability of orientation for coding continuous data dimensions.

While the use of low colour contrast is no issue in an experimental setting as used here, it may be in other situations. Moreover, we suggest to exercise caution when extrapolating our findings to displays that involve high-colour contrast features (as are often used when display clarity rather than saliency balancing is

the first requirement).

3.4.2 Colour/size interactions

We did not find any strong evidence for interactions in judgment of colour and size. A first indication for the absence of such interactions is that in single feature search it did not matter whether variation in the (task-irrelevant) second feature was present. Second, in conjunction search, manipulating colour contrast did not affect size error or vice versa. Third, no significant interaction effects were observed between factors search type (single feature, conjunction) and feature (colour, size), meaning that the difference between colour error in single feature search and conjunction search is similar to that between size error in single feature and conjunction search. All of these findings speak against interaction and hold for the analysis of the search stage fixations as well as the final decision fixations.

We should note, however, that the interaction effect for search stage fixations was marginally significant ($p = .07$) and in the same direction (biased towards colour) as we found earlier in a basic search task [Hannus *et al.*, 2006]. This could indicate that some form of interaction does actually exist during search, in line with earlier reported studies. Nevertheless, it presumably is too small to consider when designing visualisations. If interference exists, it seems to be very weak and the analysis of decision fixations showed that it does not affect the eventual search decisions. The only situation we can think of in which it potentially plays a role is a colour/size search task that requires very quick decisions.

The reader should also keep in mind that in the current study we used a limited range of (isoluminant) colours, for reasons explained earlier. Further research is needed to determine whether our finding of lack of cross-talk applies also for displays that use a larger range of saturations and which are not isoluminant (as in most actual information displays).

3.4.3 Colour/orientation interactions

Using the same criteria as above, we found two indications for interactions in judgment of colour/orientation combinations: in the search stage, colour contrast affects orientation error and orientation contrast affects colour error. This could be evidence for a symmetric interaction between colour and orientation. However, we observed some irregularities in the colour/orientation results that make us hesitate to draw any firm conclusions. We observed that increasing orientation contrast did not diminish orientation error. Also, considering that the expected value of the error was 3.3 for random fixations, we saw that orientation error during the search stage was exceptionally high (approximately 3, Figure 3.12). Apparently, saliency matching did not work properly for orientation. Further proof that orientation-trials were more difficult than size-trials is formed by the observations that only 56% of the colour/orientation but 68% of the colour/size trials met our 'final fixation criterion' (see Section 3.2.6) and that the average

number of fixations was 13.2 in a colour/size but 14.4 in a colour/orientation trial. We therefore will consider the orientation dimension in more detail below.

3.4.4 Using orientation in visualisation

Contrast matching was unsuccessful for orientation. A possible explanation can be found in [Nothdurft, 1993a]. In this study it was shown that the amount of variance in the orientation of background elements strongly affects orientation salience: increased variance in background orientation results in decreased target salience. In our stimuli, the distractor nodes as well as the links can be seen as oriented background elements. Since the stimuli contained a large number of distractors with many different orientations, background orientation variance was very high in our experiments. Contrast matching, however, was based on a search task with very little variation in background orientation (all distractors had the same orientation and there were no links). This clearly illustrates that elementary psychophysical findings cannot always be directly translated to information visualisation applications.

One might consider increasing orientation contrast to obtain better salience matches with colour and size. However, orientation contrast was already close to maximum in our stimuli (covering 63 of 180 possible angular degrees in conditions with $O=1$ and 126 in the condition with $O=2$). It is therefore impossible to further increase salience by enhancing contrast. Besides, this might result in more background variation, decreasing performance even further. While lowering the number of different orientations in the stimuli is a viable option, in practice this would also translate into reducing the amount of information that can be visualised through orientation. We therefore conclude that orientation is a less suitable feature for visualising data dimensions that cover a large variety of different values (an exception is formed by data that contain spatial dependencies, such as magnetic or flow fields; see Ware & Knight [1995] for an example).

Orientation appears to be more strongly affected by such background variation influences than are size and colour. Nevertheless, also the latter dimensions can be affected by interactions with their background. Perceived size is affected by the size of background elements, as for example demonstrated by the Ebbinghaus illusion. Colour induction may change the perceived colour of surfaces (e.g. Brenner & Cornelissen [1991]). However, such effects are different from those observed with orientation, as increased background variation tends to reduce rather than enhance such spatial interactions (e.g. Brenner & Cornelissen [2002]). This again shows that it is difficult to generalise findings from one type of display to other types and, hence, that experimentation is an important tool for optimal display design.

3.4.5 Feature hierarchy

Some authors have proposed the existence of a 'feature hierarchy' [Ebert, 2004; Healey, 2001], referring to the observation that particular features are more salient than others (or, more precisely, that variations in particular feature dimensions are more salient than variations in others). The exact details of such a hierarchy remain unclear, however. Healey seems not to make a distinction between feature interference (or interaction) and feature hierarchy; Ebert says that some features are more 'significant' perceptual cues than others. We propose to make a clear distinction between feature interference/interaction on the one hand and feature hierarchy on the other. In our view, the former refers to appearance and judgment of one feature affecting that of another, whereas feature hierarchy signifies that variations in some feature dimensions are more salient and easier to discriminate than variations in other dimensions.

Whereas Ebert seems to regard this hierarchy as static, we favour a more flexible notion. We believe that the reason why features may seem to be organised in a static hierarchy sometimes is simply that feature contrasts are ignored in such assessments. We noticed that colour contrast has to be kept very low in order to match its discriminability with that of a size contrast that is suitable for visualisation purposes (the more we increase size contrast, the more display space is occupied and the less information can be presented). As typical visualisations use highly saturated colours, it is not surprising that it is always found to be the most salient feature.

The fact that our matching procedure was effective (at least for colour and size) clearly demonstrates that the feature hierarchy is not at all static. Contrast of colour and size were successfully matched, resulting in equal performance for single feature search tasks. Our results also show that manipulation of contrasts makes it possible to match discriminability in conjunction search. This can be demonstrated by considering Figures 3.6(b) and 3.8(b): colour is more salient when $S < 1.5$; colour and size are matched when $S=1.5$; size is more salient when $S > 1.5$. Altogether, this shows that feature hierarchies are not fixed, but determined by feature contrasts. This means that the term 'hierarchy' can even be misleading in this context, as it can easily suggest that there is a fixed ordering of feature discriminability.

3.4.6 Quantification and balancing of feature contrasts in visualisation

The results of our experiments show that feature contrast and salience are closely related to each other. It appears that colour contrast determines search performance on colour but leaves size performance unaffected, and vice versa. For information visualisation these are important properties, because they allow for independent control of discriminability and salience of information that is coded by colour and size. Nonetheless, apart from some exceptions (e.g., Tufte [1986]), feature contrast appears often not to be considered an important issue in visualisation.

lisation design and most of the time designers seem to choose their contrasts by intuition. Discriminability in information dimensions of such visualisations is quite arbitrary and it is very likely that variations in some dimensions are (unpredictably) more salient than others. Nevertheless, it seems reasonable to assume that a visualisation in which the discriminability and salience of represented information directly reflects its importance is more effective than a visualisation in which there is no correspondence between these two quantities. Therefore, if we strive for visualisations with optimal effectiveness, it is necessary to quantify and carefully balance feature contrasts. Based on the above considerations we could say that existence of a clear relationship between feature contrast and salience serves as a criterion for a feature's suitability for information visualisation purposes.

For features that meet this criterion, control over salience of the information dimensions visualised by them can be obtained by first determining their contrast-salience relationships. Due to individual differences the best results would be achieved if these relationships are determined for each person separately. A drawback of our matching procedure in this respect is that it is a very time-consuming process. It would therefore be interesting to investigate whether there are simpler and more efficient ways to measure and balance an individual's sensitivity to contrasts in different features, comparable to how the classical flicker fusion test [Ives, 1912] can be used to determine psychophysical isoluminance. Such a method could then be implemented in visualisation applications and give users the opportunity to adapt such applications to their own perceptual systems. As long as such a method is not available, suboptimal results (but still better than nothing) can be expected using the average contrast sensitivities of a group of individuals.

Although outside the scope of the current experiment, another important aspect relating to quantification of visual features concerns the total number of just-noticeable differences, i.e., the number of discrete values that can accurately be coded by a feature. Some work has been done regarding this matter (e.g., Weigle *et al.* [2000] studied 'orientation categories'), but there are many unanswered questions and more research is needed on this point.

3.4.7 The use of psychophysics in visualisation

The need of perceptually motivated methods in visualisation has been recognised by many of today's visualisation researchers. The simplest approach is to incorporate facts that are already known from perception research literature. Unfortunately, as results in perception research are generally obtained by methods that do not adequately reflect visualisation practices, it is often doubtful whether they are also valid in visualisation applications. Experiments and user studies are therefore needed to verify results from perception research in more visualisation-realistic contexts (see also Kosara *et al.* [2003]). Since this is often a very time and energy consuming process, an important question is whether the benefits from

such experiments outweigh the costs.

In the work presented here, eye movements were measured during several search tasks in order to find an answer to the question whether earlier reported feature judgment interactions have any significance for information visualisation. Here a task was used that is more complex and better reflecting the kind of search tasks found in visualisation applications than those on which the previously reported interactions were based. We believe that our results are general enough to be informative for information visualisation and thereby justify the effort that was put into the experiments. We should note, however, that since in visualisation the eventual accuracy and speed with which a task is solved is usually more important than how it is solved, from a purely practical point of view it might have been sufficient to only measure the final node selection decisions and, as such, avoid the need of measuring eye movements.

As a final note, our finding that data from a simple search task could not be used to predict orientation salience in more complex tasks illustrates the risk of straightforwardly generalising research results from one domain to another and, thereby, shows the usefulness of conducting psychophysical methods in visualisation research.

3.5 Conclusion

Visual search experiments were carried out in order to find out whether earlier reported feature judgment interactions are relevant to consider in information visualisation. We specifically considered visualisations in which different data dimensions should receive equal weight and attention. Our experiments were performed with combinations of colour and size and colour and orientation. To avoid design asymmetries as well as subjects' attention being biased towards a feature with a higher salience than the others, we matched colour, orientation, and size discriminability prior to the experiments. Due to human's outstanding colour discrimination abilities, such matching inevitably requires keeping colour-contrast low or to make contrast of other features impractically or even impossibly high. We chose to use low-colour contrasts, accepting the risk that our results may not necessarily generalise to displays in which high colour contrasts are used (but that may violate other assumptions and requirements in data visualisations as well).

The most important findings from this experiment are that colour and size are features that can be used independently to represent information (at least for the range of colours that were used in our study) and that salience of features depends on the choice of feature contrasts used in a visualisation. In addition, orientation appeared to be less suitable for representing information that consists of a large range of values because it does not show a clear relationship between contrast and salience.